

AD-A058 005

CALIFORNIA UNIV BERKELEY ELECTRONICS RESEARCH LAB  
HOW TO MAKE THE LANCZOS ALGORITHM CONVERGE SLOWLY.(U)  
JAN 78 D S SCOTT

F/G 12/1

N00014-76-C-0013

UNCLASSIFIED

UCB/ERL-M78/6

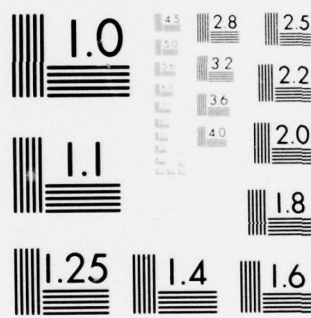
NL

| OF |

AD  
A058005



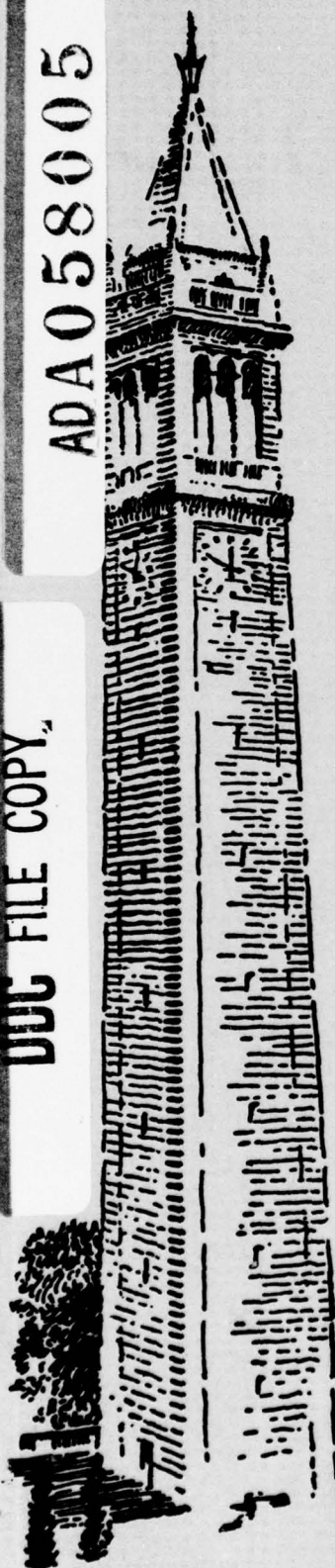
END  
DATE  
FILMED  
10-78  
DDC



MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

ADA058005

DDC FILE COPY



N/R  
LEVEL  
43



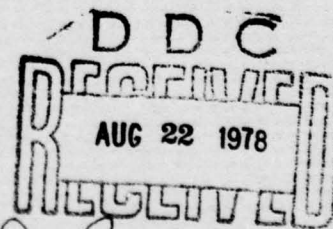
HOW TO MAKE THE LANCZOS ALGORITHM  
CONVERGE SLOWLY

by

D. S. Scott

Memorandum No. UCB/ERL M78/6

12 January 1978



**DISTRIBUTION STATEMENT A**

Approved for public release  
Distribution Unlimited

**ELECTRONICS RESEARCH LABORATORY**  
**College of Engineering**  
**University of California, Berkeley, CA 94720**

6 HOW TO MAKE THE LANCZOS ALGORITHM CONVERGE SLOWLY

by

10 D. S. Scott

9 Memorandum rept.

14  
Memorandum No. UCB/ERL-M78/6

11 12 January 1978

12 18p

15 N00014-76-C-0413

ELECTRONICS RESEARCH LABORATORY

College of Engineering  
University of California, Berkeley  
94720

127 550 78 07 27 016

# HOW TO MAKE THE LANCZOS ALGORITHM CONVERGE SLOWLY

D.S. Scott\*

## Abstract

The Paige style Lanczos algorithm is an iterative method for finding a few eigenvalues of large sparse symmetric matrices. Some beautiful relationships among the elements of the eigenvectors of a symmetric tridiagonal matrix are used to derive a perverse starting vector which delays convergence as long as possible. Why such slow convergence is never seen in practice is also examined.

Key words: eigenvalues, Lanczos algorithm, sparse symmetric matrices

ACCESSION FOR	
NTIS	White Section <input checked="" type="checkbox"/>
DDC	Bull Section <input type="checkbox"/>
UNANNOUNCED	<input type="checkbox"/>
JUSTIFICATION	
<i>letter on file</i>	
BY	
DISTRIBUTION/AVAILABILITY CODES	
Doc	ATAC
A	

\* Mathematics Department and Electronics Research Laboratory, University of California, Berkeley, CA 94720. The work of the author was supported by Office of Naval Research Contract N00014-76-C-0013.



### 1. Introduction

In 1950 Lanczos [2] presented an algorithm for reducing a symmetric matrix, call it  $A$ , to tridiagonal form. The algorithm begins with an arbitrary unit vector  $q_1$ . It produces a tridiagonal matrix  $T$  and an orthogonal matrix  $Q$  such that

$$(1) \quad Q^* A Q = T \quad \text{and} \quad Q e_1 = q_1.$$

In practice the algorithm could not compete in speed or accuracy with later methods based on explicit orthogonal transformations.

In 1971 Paige [3] introduced a modified version of the algorithm which could be used effectively to find a few eigenvalues, and their eigenvectors too if desired, of a large sparse symmetric matrix. Paige suggested terminating the process prematurely at, say, the  $j^{\text{th}}$  step with  $T_j$  the  $j \times j$  leading principal minor of  $T$  and  $Q_j$  the first  $j$  columns of  $Q$  in hand. In exact arithmetic

$$(2) \quad Q_j^* A Q_j = T_j.$$

Let the spectral decomposition of  $T_j$  be

$$(3) \quad T_j = S_j \Theta S_j^* \quad \text{with} \quad \Theta = \text{diag}(\theta_1^j, \theta_2^j, \dots, \theta_j^j) \quad \text{and} \quad S_j^* S_j = I_j.$$

Define

$$Y_j = (y_1^j, y_2^j, \dots, y_j^j) = Q_j S_j.$$

Then  $\theta_1^j, \theta_2^j, \dots, \theta_j^j$  are the Rayleigh-Ritz approximations to the eigenvalues of  $A$  (commonly called Ritz values) derivable from the subspace spanned by the columns of  $Q_j$ , and  $y_1^j, y_2^j, \dots, y_j^j$  are the corresponding Ritz vectors. The norm of the residual of  $y_i^j$ , namely  $\|A y_i^j - \theta_i^j y_i^j\|$ , is a bound on the accuracy of  $\theta_i^j$  as an approximation to an eigenvalue of  $A$ .

The Kaniel-Paige error estimates [1,3] lead us to expect that some of the eigenvalues of  $T_j$  should converge (have negligible error bound) for  $j \ll n$ , provided only that the starting vector  $q_1$  is not pathologically deficient in the corresponding eigendirections of  $A$ . Numerical tests by Paige and other researchers have confirmed that convergence occurs relatively quickly. Despite this abundance of evidence, Paige was unable to prove that convergence of some Ritz value must occur before  $j = n = \dim(A)$  at which point, in exact arithmetic,  $T$  is similar to  $A$  so that all the "approximations" are exact and all the bounds are zero.

There are several interesting unresolved problems connected with the Lanczos process. Except in its last section, this paper is restricted to the theoretical behavior of the algorithm in the context of exact arithmetic. In the following section we derive some beautiful relationships among the elements of the eigenvectors of a symmetric tridiagonal matrix which may be of interest in their own right. In Section 3 these results are applied to obtain formulas for the Lanczos starting vector. In Section 4 these formulas are used to find a perverse starting vector for matrices with well separated eigenvalues which delays convergence until  $j = n$ . Section 5 generalizes the construction to matrices with close or multiple eigenvalues to yield a vector which delays convergence for a long time. The final section will indicate why such slow convergence is never seen in practice.

## 2. The Eigenvectors of a Symmetric Tridiagonal Matrix

Definition. Let  $\text{adj}(R)$  be the transpose of the matrix of cofactors of  $R$ . This is usually called the adjugate or classical adjoint of  $R$ . By the Cauchy-Binet theorem

$$(1) \quad R \text{adj}(R) = \det(R)I.$$

Example.

$$R = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 2 & -1 & 1 \end{bmatrix}$$

$$\text{adj}(R) = \begin{bmatrix} 2 & -1 & -1 \\ 1 & -1 & 0 \\ -3 & 1 & 1 \end{bmatrix}$$

$$R \text{ adj}(R) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

Theorem 2.1 (Thompson and McEntegert, 1968). Let  $A = Z\Lambda Z^*$  with  $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ ,  $Z = (z_1, z_2, \dots, z_n)$  and  $Z^*Z = I$ . Then for  $i = 1, 2, 3, \dots, n$

$$\text{adj}(\lambda_i I - A) = \prod_{\substack{j=1 \\ j \neq i}}^n (\lambda_j - \lambda_i) z_i z_i^* = \chi_A'(\lambda_i) z_i z_i^*,$$

where  $\chi_A'(\mu)$  is the derivative of the characteristic polynomial of  $A$ .

Note that if  $\lambda_i$  is a multiple eigenvalue of  $A$ , then  $\chi_A'(\lambda_i) = 0$ , so that the ambiguity in the choice of eigenvectors doesn't matter.

Proof. Let  $\mu \neq \lambda_i$ , for all  $i$ , so that  $(\mu I - A)^{-1}$  exists. Then

$$\begin{aligned} \text{adj}(\mu I - A) &= \det(\mu I - A) (\mu I - A)^{-1} \\ &= \chi_A(\mu) Z (\mu I - \Lambda)^{-1} Z^* \\ &= Z \Delta Z^* \end{aligned}$$

where  $\Delta$  is diagonal and  $\Delta_{kk} = \frac{\chi_A(\mu)}{\mu - \lambda_k} = \prod_{\substack{j=1 \\ j \neq k}}^n (\mu - \lambda_j)$ . Since computing cofactors



does not involve division,  $\text{adj}(R)$  is a continuous function of  $R$ . Therefore the last equation must hold even for  $\mu = \lambda_i$ . Setting  $\mu = \lambda_i$ , for  $i = 1, 2, 3, \dots, n$ , yields

$$\text{adj}(\lambda_i I - A) = Z \Delta Z^*,$$

where

$$\Delta_{kk} = \begin{cases} 0 & \text{if } k \neq i, \\ \prod_{\substack{j=1 \\ j \neq i}}^n (\lambda_i - \lambda_j) & \text{if } k = i. \end{cases}$$

Since  $\prod_{\substack{j=1 \\ j \neq i}}^n (\lambda_i - \lambda_j) = \chi'_A(\lambda_i)$ , the result follows.  $\square$

Thompson and McEntegert were working with general Hermitian matrices. The application of their theorem to tridiagonal matrices was made by Paige [3].

Notation. Let

$$T_{r,t} \equiv \begin{bmatrix} \alpha_r & \beta_r & & & & \\ \beta_r & \alpha_{r+1} & \beta_{r+1} & & & \\ & \beta_{r+1} & \alpha_{r+2} & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & \beta_{t-1} \\ \bigcirc & & & & \beta_{t-1} & \alpha_t \end{bmatrix},$$

Let  $\chi_{r,t}(\mu) \equiv \det(\mu I - T_{r,t})$ , the characteristic polynomial of  $T_{r,t}$ , and let  $\chi_{r,r-1}(t) \equiv 1$  for all  $r$ .

Theorem 2.2 (Paige, 1971). Let  $T = T_{1,n} = S\Theta S^*$  with  $\Theta = \text{diag}(\theta_1, \theta_2, \dots, \theta_n)$  and  $S^*S = I$ . Then for  $r \leq t$  and all  $i$ ,

$$x'_{1,n}(\theta_i) s_{ri} s_{ti} = x_{1,r-1}(\theta_i) \beta_r \beta_{r+1} \cdots \beta_{t-1} x_{t+1,n}(\theta_i).$$

In particular

$$x'_{1,n}(\theta_i) s_{ri}^2 = x_{1,r-1}(\theta_i) x_{r+1,n}(\theta_i).$$

Proof. By Theorem 2.1

$$(2) \quad \text{adj}(\theta_i I - T) = x'_{1,n}(\theta_i) s_i s_i^*.$$

The  $(r,t)$  element of the RHS of (2) is  $x'_{1,n}(\theta_i) s_{ri} s_{ti}$ . Because of the tridiagonal form of  $T$ , the  $(r,t)$  element of the LHS of (2) is  $x_{1,r-1}(\theta_i) \beta_r \beta_{r+1} \cdots \beta_{t-1} x_{t+1,n}(\theta_i)$ . For example,

$$\theta_i I - T_{1,6} = \begin{bmatrix} \textcircled{\theta_i - \alpha_1} & \beta_1 & & & & \\ -\beta_1 & \theta_i - \alpha_2 & -\beta_2 & & & \\ & \textcircled{-\beta_2} & \theta_i - \alpha_3 & -\beta_3 & & \\ & & -\beta_3 & \theta_i - \alpha_4 & -\beta_4 & \\ & & & -\beta_4 & \theta_i - \alpha_5 & -\beta_5 \\ & & & & -\beta_5 & \theta_i - \alpha_6 \end{bmatrix}$$

The circled elements contribute to the  $(2,3)$  cofactor. Note that the minus signs on the  $\beta$ 's cancel with the alternating signs associated with the cofactors. □

### 3. Formulas for Starting Vectors

The Lanczos algorithm begins with an arbitrary unit vector  $q_1$  and terminates with a tridiagonal matrix  $T$  and an orthogonal matrix  $Q$  such

that  $q_1$  is the first column of  $Q$  and  $AQ = QT$ . The process is geometric, i.e. it is invariant under orthogonal changes in coordinates. The coordinates which give the most insight into the process are the eigenvectors of  $A$ .

In these coordinates the operator  $A$  is diagonal and the matrix  $Q$  becomes the transpose of  $S$ , the matrix of eigenvectors of  $T$ . The matrix equation  $AQ = QT$  becomes

$$(1) \quad \Lambda S^* = S^* T$$

where  $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ .

**Theorem 3.1.** Let  $\Lambda S^* = S^* T$  as above. Then for  $i = 1, 2, \dots, n$

- (i)  $s_{1i} s_{ni} x'_\Lambda(\lambda_i) = \beta_1 \beta_2 \cdots \beta_{n-1} \equiv \pi_n$  a constant,
- (ii)  $s_{ni}^2 x'_\Lambda(\lambda_i) = x_{1,n-1}(\lambda_i)$ .

Proof. Since  $\Lambda$  is similar to  $T$ ,  $x_\Lambda(\mu) = x_T(\mu) = x(\mu)$ .

(i) This is Theorem 2.2 with  $r = 1$  and  $t = n$ .

(ii) This is Theorem 2.2 with  $r = n$  and  $t = n$ . □

In order to refer to the Lanczos vectors, we need names for the columns of  $S^*$ . For this purpose, when  $A = \Lambda$ , we define

$$(2) \quad P \equiv (p_1, p_2, \dots, p_n)^T \equiv S^*.$$

Theorem 3.1(i) relates the first Lanczos vector  $p_1$  to the last Lanczos vector  $p_n$ . Theorem 3.1(ii) relates  $p_n$  to the eigenvalues of  $T_{1,n-1}$  which are the approximations to eigenvalues of  $\Lambda$  furnished by the  $(n-1)^{\text{st}}$  step of the Lanczos algorithm. Since  $T_{1,n}$  is similar to  $A$ , the Cauchy Interlace theorem requires that the eigenvalues of  $T_{n-1}$ , call them  $\mu_1, \mu_2, \dots, \mu_{n-1}$ ,

satisfy the inequalities

$$(3) \quad \lambda_1 \leq \mu_1 \leq \lambda_2 \leq \dots \leq \mu_{n-1} \leq \lambda_n.$$

Since  $T_{1,n}$ , the tridiagonal matrix produced by the Lanczos algorithm, will be unreduced, the interlacing must be strict.

Theorem 3.2. Let  $A = Z\Lambda Z^*$  with  $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$  and  $Z^*Z = I$ . Then the Lanczos algorithm run with a starting vector  $q_1$  produces a  $T_{1,n-1}$  with eigenvalues  $\mu_1 < \mu_2 < \dots < \mu_{n-1}$  if and only if  $q_1 = Zp_1$ , where

$$p_{i1}^2 = \pi_n^2 \left[ \prod_{\substack{j=1 \\ j \neq i}}^n (\lambda_i - \lambda_j) \prod_{j=1}^{n-1} (\lambda_i - \mu_j) \right]^{-1} = \pi_n^2 [X'(\lambda_i) X_{1,n-1}(\lambda_i)]^{-1}.$$

Strict interlacing is required to make the quantity in the brackets positive for all  $i$ .

Proof. The Lanczos algorithm produces the same  $T$  whether it runs on the pair  $(A, q_1)$  or  $(\Lambda, p_1)$ . Combining the two parts of Theorem 3.1 and changing to the  $P$  notation yields

$$(4) \quad p_{i1}^2 X'(\lambda_i) X_{1,n-1}(\lambda_i) = \pi_n^2,$$

for any starting vector  $p_1$ .

If  $p_1$  is known, then by interpolation,  $x_{1,n-1}(\mu)$  can be found from (4), up to a constant factor. Hence  $\mu_1, \mu_2, \dots, \mu_{n-1}$ , the zeros of  $x_{1,n-1}(\mu)$  can be found.

If  $\mu_1, \mu_2, \dots, \mu_{n-1}$  are given, then  $p_{i1}^2$ , for  $i = 1, 2, \dots, n$ , can be found from (4) up to the multiplicative factor  $\pi_n^2$  which can be determined



by the required normalization of  $p_1$ . The ambiguity in the choice of sign for each component of  $p_1$  merely reflects the choice of sign for each eigenvector of  $T$ . All choices yield the same tridiagonal matrix  $T$ .  $\square$

The required  $q_1$  depends on both the eigenvalues and on the eigenvectors of  $A$ . The expression  $q_1 = Zp_1$  clarifies their roles;  $Z$  is independent of the  $\lambda_i$ , while  $p_1$  is independent of  $Z$ .

Example. Let  $\Lambda = \text{diag}(1,3,5,7,9)$  and  $\mu_i = 2i$ , for  $i = 1,2,3,4$ .

$$\chi'(1)\chi_\mu(1) = (1-3)(1-5)(1-7)(1-9)(1-2)(1-4)(1-6)(1-8) = 40320$$

$$\chi'(3)\chi_\mu(3) = (3-1)(3-5)(3-7)(3-9)(3-2)(3-4)(3-6)(3-8) = 1440$$

$$\chi'(5)\chi_\mu(5) = (5-1)(5-3)(5-7)(5-9)(5-2)(5-4)(5-6)(5-8) = 576$$

$$\chi'(7)\chi_\mu(7) = (7-1)(7-3)(7-5)(7-9)(7-2)(7-4)(7-6)(7-8) = 1440$$

$$\chi'(9)\chi_\mu(9) = (9-1)(9-3)(9-5)(9-7)(9-2)(9-4)(9-6)(9-8) = 40320$$

so

$$p_{11} = p_{51} = \pi_5 / \sqrt{40320} = .00498\pi_5$$

$$p_{21} = p_{41} = \pi_5 / \sqrt{1440} = .02635\pi_5$$

$$p_{31} = \pi_5 / \sqrt{576} = .04167\pi_5$$

By normalization  $\pi_5 = 17.749$  and

$$p_1 = (.0880, .4677, .7396, .4677, .0880)^*.$$

The Lanczos algorithm run on  $\Lambda$  with  $p_1$  as the starting vector yielded a  $T_4$  with eigenvalues 2, 4, 6, and 8 correct to the precision of the machine used.  $\square$



4. Slow Convergence

Before examining the convergence properties of the Lanczos algorithm in the light of Theorem 3.2, it is useful to describe in more detail the properties of the Rayleigh-Ritz procedure. Let  $W$  be any subspace of  $\mathbb{R}^n$  and let  $P_W$  denote the orthogonal projection of  $\mathbb{R}^n$  onto  $W$ . Then the Rayleigh-Ritz approximations to eigenpairs of  $A$  obtained from  $W$  are precisely the eigenpairs of  $P_W A$  whose eigenvectors lie in  $W$ .

Theorem 4.1. If  $V \subseteq W$  then the Rayleigh-Ritz approximations for  $A$  obtained from  $V$  are the same as the Rayleigh-Ritz approximations for  $P_W A$  obtained from  $V$ . Further if  $(y, \theta)$  is a Ritz pair then  $\|(P_W A)y - y\theta\| \leq \|Ay - y\theta\|$ , with equality holding if and only if the residual vector,  $Ay - y\theta$ , lies in  $W$ .

Proof. Let  $P_V$  be the orthogonal projection onto  $V$ . Then  $P_V(P_W A) = (P_V P_W)A = P_V A$ , since  $V \subseteq W$ . Since  $y \in V \subseteq W$ ,  $\|(P_W A)y - y\theta\| = \|P_W(Ay - y\theta)\|$ . Finally since  $P_W$  is an orthogonal projection,  $\|P_W(Ay - y\theta)\| \leq \|Ay - y\theta\|$ , with equality holding if and only if  $Ay - y\theta \in W$ .  $\square$

Corollary. If  $V$  and  $W$  are nested Krylov subspaces of different dimensions, then  $\|Ay - y\theta\| = \|P_W Ay - y\theta\|$ .

Proof. If  $V = K_j(q_1)$  and  $W = K_k(q_1)$  for  $k > j$ , then, since  $y \in K_j(q_1)$ ,  $Ay - y\theta \in K_{j+1}(q_1) \subseteq K_k(q_1) = W$ .  $\square$

We now examine the convergence properties of the Lanczos algorithm. The reader is directed to Section 1 for the terminology.

Theorem 4.2. Suppose that the Lanczos algorithm when run on  $(A, q_1)$  produces  $\mu_1, \mu_2, \dots, \mu_{n-1}$  as Ritz values at the  $n-1^{\text{st}}$  step. Let  $(y, \theta)$  be a Ritz pair from any step except the  $n^{\text{th}}$ . Then

$$\gamma \equiv \|Ay - y\theta\| \geq \delta_\mu / 2 ,$$

$$\text{where } \delta_\mu = \min_{\substack{i \leq n-1 \\ k \leq n}} |\mu_i - \lambda_k| .$$

Proof. For any vector  $x$  and any scalar  $\tau$  it is well known that

$$(5) \quad \min_{i \leq n} |\lambda_i - \tau| \leq \|Ax - x\tau\| .$$

In particular,

$$(6) \quad \min_{i \leq n} |\lambda_i - \theta| \leq \|Ay - y\theta\| = \gamma , \text{ and}$$

$$(7) \quad \min |\mu_i - \theta| \leq \|P_W Ay - y\theta\| = \gamma ,$$

with the last equality following from the Corollary. The smallest  $\gamma$  which can satisfy both (6) and (7) is  $\delta_\mu / 2$ .  $\square$

In practice  $\gamma$  can not be as small as  $\delta_\mu / 2$ . However no significantly stronger bound can be obtained. In particular, the smallest residual at the  $n-1^{\text{st}}$  step can be  $\delta_\mu$ .

The combination of Theorem 3.2 and Theorem 4.2 yields the following.

Theorem 4.3. Let  $A$  be a symmetric matrix with eigenvalues

$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ . Let  $\delta_A \equiv \min_{i \neq k} |\lambda_i - \lambda_k|$ . Then there exists a starting vector for the Lanczos algorithm such that the residual norm of any Ritz vector at any step  $j < n$  will be larger than  $\delta_A / 4$ .

Proof. If  $A$  has multiple eigenvalues then  $\delta_A = 0$  and any vector will do, so we may assume that  $A$  has distinct eigenvalues. Let  $\mu_i = (\lambda_i + \lambda_{i+1})/2$ , for  $i = 1, 2, \dots, n-1$ , and let  $q_1$  be any starting vector generated by Theorem 3.2. With this choice of  $q_1$ ,  $\mu_1, \mu_2, \dots, \mu_{n-1}$  will be the Ritz values at the  $n-1^{\text{st}}$  step and  $\delta_\mu = \delta_A/2$ . The result now follows from Theorem 4.2.  $\square$

If the spectrum of  $A$  is such that  $\delta_A/4$  is larger than some given convergence tolerance, then Theorem 4.3 shows that there exist perverse starting vectors which delay convergence until the  $n^{\text{th}}$  step. This result does not imply that no earlier Ritz value is accurate enough, it only guarantees that the corresponding bound will not reveal such accuracy. In the previous example of  $\Lambda = \text{diag}(1, 3, 5, 7, 9)$  and  $\mu_i = 2i$ , for  $i = 1, 2, 3, 4$ ,  $\theta_2^3$ , the middle eigenvalue of  $T_{1,3}$  is 5, correct to working accuracy. The corresponding bound is 1.25, which shows that this fortuitous accuracy is due to the symmetry of the example, rather than the accuracy of the Ritz vector.

## 5. The Problem of Clustered Eigenvalues

If the spectrum of  $A$  is such that  $\delta_A/4$  is smaller than the given convergence tolerance, Theorem 4.3 does not guarantee slow convergence. However starting vectors can still be found which delay convergence a long time.

Theorem 5.1. Let  $W$  be an  $A$ -invariant subspace of maximal dimension such that  $\bar{A} \equiv A$  restricted to  $W$  is such that  $\delta_{\bar{A}}/4$  is larger than the given convergence tolerance. Let  $m = \dim W$ . Then there exists a starting vector for  $A$  which delays convergence until the  $m^{\text{th}}$  step.

Proof. Apply Theorem 4.3 to  $\bar{A}$  to yield a starting vector  $q_1$ . Since the Lanczos algorithm run on  $(A, q_1)$  yields the same  $T$  as the Lanczos algorithm run on  $(\bar{A}, q_1)$ , this  $q_1$  will do.  $\square$

In general, it may be possible to delay convergence even longer.

## 6. The Beneficial Effects of Rounding Errors

The slow convergence discussed in the previous sections never seems to occur in practice. The reason for this lies in the formula for  $p_1$  given in Theorem 3.2,

$$(1) \quad p_{i1}^2 = \pi_n^2 [X'(\lambda_i) X_{1,n-1}(\lambda_i)]^{-1}.$$

First we give three examples.

Example 1. Linear distribution.

$$\begin{aligned} \lambda_i &= i & \text{for } i = 1, 2, \dots, 50 \\ \mu_i &= \frac{(\lambda_i + \lambda_{i+1})}{2} = i + \frac{1}{2} & \text{for } i = 1, 2, \dots, 49. \end{aligned}$$

$p_1$  was computed by Theorem 3.2.  $p_1$  is symmetric from top to bottom. The largest elements of  $p_1$  are  $p_{25,1} = p_{26,1} = .397$ . The smallest elements of  $p_1$  are  $p_{1,1} = p_{50,1} = .25 \times 10^{-14}$ .



Example 2. Geometric distribution.

$$\lambda_i = (1.1)^i \quad \text{for } i = 1, 2, \dots, 50$$

$$\mu_i = \frac{(\lambda_i + \lambda_{i+1})}{2} \quad \text{for } i = 1, 2, \dots, 49$$

The largest element of  $p_1$  is  $p_{8,1} = .495$ . The smallest element of  $p_1$  is  $p_{50,1} = .162 \times 10^{-52}$ .

Example 3. Tchebychev distribution.

$$\lambda_i = \cos\left(\frac{i\pi}{51}\right) \quad \text{for } i = 1, 2, \dots, 50$$

$$\mu_i = \frac{(\lambda_i + \lambda_{i+1})}{2} \quad \text{for } i = 1, 2, \dots, 49$$

The largest elements of  $p_1$  are  $p_{25,1} = p_{26,1} = .228$ . The smallest elements of  $p_1$  are  $p_{1,1} = p_{50,1} = .642 \times 10^{-3}$ .

The tiny elements of  $p_1$  in Examples 1 and 2 are due to the large variation in magnitudes of the numbers  $\{\chi'(\lambda_k) | k=1, 2, \dots, n\}$ . Most practical examples also show large variations, which leads to a perverse starting vector with some tiny eigenvectors. Such tiny components are unlikely to appear in a randomly chosen vector. More importantly, such tiny components are unstable in the face of rounding errors.

By standard rounding error analysis

$$(2) \quad q_1 = Zp_1 + f, \quad \|f\| \leq n^{3/2}e$$

where  $e$  is the relative machine precision. In exact arithmetic,

$$(3) \quad Z^*q_1 = p_1 + Z^*f.$$



Unless  $Z$  has some special symmetry, the term  $Z^*f$  will swamp any tiny component of  $p_1$ . That is,  $q_1$  is unlikely to have any eigencomponents much smaller than  $\epsilon$ . Even if  $q_1$  had precisely the eigencomponents desired, the first step of the Lanczos algorithm would obliterate the small components unless the example were specially rigged. The first step of the algorithm computes  $q_2$  as

$$(4) \quad \beta_1 q_2 = (A - \alpha_1 I)q_1 + f, \quad \text{where } \|f\| \leq n^{3/2} \epsilon \|A\|.$$

Again  $f$  will be randomly distributed among the various eigendirections and will prevent  $q_2$  from inheriting any tiny components from  $q_1$ .

We have been able to observe delayed convergence for large matrices only for two classes. Tchebychev distributions come very close to minimizing the variation in  $\chi'_{1,n}(\lambda_i)$ . Therefore Tchebychev distributions (even on fairly large problems) do not have components of  $p_1$  smaller than  $\epsilon$ . The other class of examples is diagonal matrices in which the rounding errors are uncoupled with tiny elements of  $p_1$  (which is  $q_1$ ) are not swamped by small multiples of much larger elements.

References

- [1] Kaniel, S., "Estimates for Some Computational Techniques in Linear Algebra," Math. Comp. v.20, pp. 369-378, 1966.
- [2] Lanczos, C., "An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators," J. Res. Nat. Bur. Standards, v.45, pp. 255-282, 1950.
- [3] Paige, C.C., The Computation of Eigenvalues and Eigenvectors of Very Large Sparse Matrices, Ph.D. Thesis, University of London, 1971.
- [4] Thompson, R.C. and McEntegert, P., "Principal Submatrices II: The Upper and Lower Quadratic Inequalities," Linear Algebra and Its Applications, v.1, pp. 211-243, 1968.